

# ON THE ENTROPY OF A FAMILY OF RANDOM SUBSTITUTIONS

JOHAN NILSSON

jnilsson@math.uni-bielefeld.de

18th January 2013

## Abstract

The generalised random Fibonacci chain is a stochastic extension of the classical Fibonacci substitution and is defined as the rule mapping  $0 \mapsto 1$  and  $1 \mapsto 1^i 0 1^{m-i}$  with probability  $p_i$ , where  $p_i \geq 0$  with  $\sum_{i=0}^m p_i = 1$ , and where the random rule is applied each time it acts on a 1. We show that the topological entropy of this object is given by the growth rate of the set of inflated generalised random Fibonacci words.

2010 Mathematics Subject Classification: 68R15 Combinatorics on words, 05A16 Asymptotic enumeration, 37B10 Symbolic dynamics.

## 1 INTRODUCTION

Meyer sets form an important class of point sets that can be considered as generalisations of lattices [8, 9]. They have recently been studied in connection with mathematical models of quasicrystals, see [10, 6, 2] and references therein. One of their common features is the existence of a non-trivial point spectrum. It shows up as a relatively dense set of point measures in their diffraction measure [12], even under the constraint that one only considers point measures whose intensity is at least a given positive fraction of the central intensity.

The majority of papers so far has concentrated on deterministic Meyer sets, such as those obtained from Pisot inflations or from the projection method. All these examples have zero entropy. On the other hand, it is well known that Meyer sets with entropy exist (such as the set  $2\mathbb{Z}$  combined with an arbitrary subset of  $2\mathbb{Z} + 1$ : compare [3]) but relatively little is known about them. Despite having entropy, the spectral result of Strungaru [12] still applies, and the point spectrum is non-trivial.

One possibility to define special classes of Meyer sets with entropy is via random inflation rules, as proposed in [5]. Here we follow their line of approach for a family of generalised Fibonacci substitutions, parametrised by a number  $m \in \mathbb{N}$ . The spectral nature for  $m = 1$  was studied in [5, 7]. Here we reconsider this case and its generalisation from the point of

view of entropy. In fact we prove that the topological entropy can simply be calculated by a suitable (and easily accessible) subset of sub-words (or factors), thus proving an (implicit) conjecture from [5] and its generalisation to the entire family.

Let us introduce the generalised random Fibonacci chain by the generalised substitution

$$\theta : \begin{cases} 0 \mapsto 1 \\ 1 \mapsto \begin{cases} 01^m & \text{with probability } p_0 \\ 101^{m-1} & \text{with probability } p_1 \\ \vdots \\ 1^m 0 & \text{with probability } p_m \end{cases} \end{cases} \quad (1)$$

where  $p_i \geq 0$  and  $\sum_{i=0}^m p_i = 1$  and where the random rule is applied each time  $\theta$  acts on a 1.

In [5] Godrèche and Luck define the random Fibonacci chain by the generalised substitution given by  $\theta$ , from (1), in the special case  $m = 1$ . They introduce the random Fibonacci chain when studying quasi-crystalline structures and tilings in the plane. In their paper it is claimed without proof that the topological entropy of the random Fibonacci chain is given by the growth rate of the set of inflated random Fibonacci words. In [11] this fact was proven, and we shall here give a proof of a more general result when we consider the generalised random Fibonacci substitution given by  $\theta$  from (1).

Before we can state our main theorem, we need to introduce some notation. A word  $w$  over an alphabet  $\Sigma$  is a finite sequence  $w_1 w_2 \dots w_n$  of symbols from  $\Sigma$ . We let here  $\Sigma = \{0, 1\}$ . We denote a sub-word or a factor of  $w$  by  $w[a, b] = w_a w_{a+1} w_{a+2} \dots w_{b-1} w_b$  and similarly for a set of words we let  $W[a, b] = \{w[a, b] : w \in W\}$ . By  $|\cdot|$  we mean the length of a word and the cardinality of a set. Note that  $|w[a, b]| = b - a + 1$ .

For two words  $u = u_1 u_2 u_3 \dots u_n$  and  $v = v_1 v_2 v_3 \dots v_m$  we denote by  $uv$  the concatenation of the two words, that is,  $uv = u_1 u_2 u_3 \dots u_n v_1 v_2 \dots v_m$ . Similarly we let for two sets of words  $U$  and  $V$  their product be the set  $UV = \{uv : u \in U, v \in V\}$  containing all possible concatenations.

Letting  $\theta$  act on the word 0 repeatedly yields an infinite sequence of words  $r_n = \theta^{n-1}(0)$ . We know that  $r_1 = 0$  and  $r_2 = 1$ . But  $r_3$  is one of the words  $1^i 0 1^{m-i}$  for  $0 \leq i \leq m$  with probability  $p_i$ . The sequence  $\{r_n\}_{n=1}^\infty$  converges in distribution to an infinite random word  $r$ . We say that  $r_n$  is an inflated word (under  $\theta$ ) in generation  $n$  and we introduce here sets that correspond to all inflated words in generation  $n$ ;

**Definition 1.** For a fixed  $m \geq 1$  let  $A_0 := \emptyset$ ,  $A_1 := \{0\}$  and  $A_2 := \{1\}$  and for  $n \geq 3$  define recursively

$$A_n = \bigcup_{i=0}^m \prod_{j=0}^m A_{n-1-\delta_{ij}}, \quad (2)$$

where  $\delta$  is the Kronecker symbol, that is  $\delta_{ij} = 1$  if  $i = j$  and 0 otherwise and we let  $A := \lim_{n \rightarrow \infty} A_n$ .

We shall later on, as a direct consequence of Proposition 7, see that the set  $A$  is a well-defined set. The integer  $m \geq 1$  is arbitrary, but fixed during the whole process. Therefore we choose to not have  $m$  as a parameter when denoting a set of inflated generalised random Fibonacci words,  $A_n$ .

It is clear from the definition of  $A_n$  that all elements in  $A_n$  have the same length, that is, for all  $x, y \in A_n$  we have  $|x| = |y|$ . We shall frequently use the length of the elements in  $A_n$ , and therefore we introduce the notation

$$l_n := |x| \quad \text{for } x \in A_n, \quad n \geq 1.$$

From the recursion (2) and the definition of  $l_n$  we have immediately the following proposition

**Proposition 2.** The numbers  $l_n$  fulfil the recursion relation,  $l_1 = 1$ ,  $l_2 = 1$  and  $l_n = m \cdot l_{n-1} + l_{n-2}$  for  $n \geq 3$ .

For a word  $w$  we say that  $x$  is a *factor* of  $w$  if there are two words  $u, v$  such that  $w = uxv$ . The factor set  $F(S, n)$ , of a set of words  $S$ , is the set of all factors of length  $n$  of the words in  $S$ . We introduce the abbreviated notation  $F_n$  for the special factor sets of the generalised random Fibonacci words,

$$F_n := F(A, l_n).$$

It seems to be a hard problem to give an explicit expression for the size of  $|F_n|$ . We have to leave that question open, but we shall give a rough upper bound of  $|F_n|$ .

The topological entropy of the generalised random Fibonacci chain is defined as the limit  $\lim_{n \rightarrow \infty} \frac{1}{n} \log |F(A, n)|$ . The existence of this limit is direct by Fekete's lemma [4] since we have sub-additivity,  $\log |F(A, n+p)| \leq \log |F(A, n)| + \log |F(A, p)|$ . We can now state the main result in this paper.

**Theorem 3.** The logarithm of the growth rate of the size of the set of inflated generalised random Fibonacci words equals the topological entropy of the generalised random Fibonacci sequence, that is

$$\lim_{n \rightarrow \infty} \frac{\log |A_n|}{l_n} = \lim_{n \rightarrow \infty} \frac{\log |F_n|}{l_n}. \quad (3)$$

The outline of the paper is that we start by studying the set  $A_n$  by looking at sets of prefixes of the set  $A_n$ . Next we give a finite method for finding the factor set  $F_n$  and finally we present an estimate of  $|F_n|$  in terms of  $|A_n|$ , leading to the proof of Theorem 3.

## 2 SETS OF INFLATED WORDS

In this section we shall study the sets of inflated generalised Fibonacci words, the  $A_n$  sets. The main results of the section give how to calculate the size of  $A_n$ . The underlying idea in proving our results here is to rewrite and consider the union (2) as a union of two sets only. We obtain this by collecting all but the last union term in (2) into one set and similarly all but the first one into a second set. We can then factor out an  $A_n$ -term from these sets, and thereby apply an induction argument on the achieved smaller set.

The first result of this section gives how to simplify the overlap we obtain when rewriting and generalising (2) in the manner just described above.

**Proposition 4.** *For a fixed  $m \geq 1$  we have for  $k \geq 1$  and  $n \geq 2$*

$$\begin{aligned} \left( A_n \left( \bigcup_{i=0}^k \prod_{j=0}^k A_{n-\delta_{ij}} \right) \right) \cap \left( \left( \bigcup_{i=0}^k \prod_{j=0}^k A_{n-\delta_{ij}} \right) A_n \right) = \\ = A_n \left( \bigcup_{i=0}^{k-1} \prod_{j=0}^{k-1} A_{n-\delta_{ij}} \right) A_n. \quad (4) \end{aligned}$$

*Proof.* It is clear that

$$A_n \left( \bigcup_{i=0}^k \prod_{j=0}^k A_{n-\delta_{ij}} \right) \supseteq A_n \left( \bigcup_{i=0}^{k-1} \prod_{j=0}^k A_{n-\delta_{ij}} \right) = A_n \left( \bigcup_{i=0}^{k-1} \prod_{j=0}^{k-1} A_{n-\delta_{ij}} \right) A_n$$

and symmetrically

$$\left( \bigcup_{i=0}^k \prod_{j=0}^k A_{n-\delta_{ij}} \right) A_n \supseteq A_n \left( \bigcup_{i=0}^{k-1} \prod_{j=0}^{k-1} A_{n-\delta_{ij}} \right) A_n.$$

This gives the inclusion

$$\left( A_n \left( \bigcup_{i=0}^k \prod_{j=0}^k A_{n-\delta_{ij}} \right) \right) \cap \left( \left( \bigcup_{i=0}^k \prod_{j=0}^k A_{n-\delta_{ij}} \right) A_n \right) \supseteq$$

$$\supseteq A_n \left( \bigcup_{i=0}^{k-1} \prod_{j=0}^{k-1} A_{n-\delta_{ij}} \right) A_n. \quad (5)$$

To prove that we actually have equality in (5); assume for contradiction that the inclusion in (5) is strict. Then there is an element  $x$  in the left hand side intersection which is not in the right hand side set of (5). The word  $x$  is then an element in

$$\left( A_n \left( \bigcup_{i=0}^k \prod_{j=0}^k A_{n-\delta_{ij}} \right) \right) \setminus \left( A_n \left( \bigcup_{i=0}^{k-1} \prod_{j=0}^{k-1} A_{n-\delta_{ij}} \right) A_n \right) \quad (6)$$

and then also an element in the set

$$\left( \left( \bigcup_{i=0}^k \prod_{j=0}^k A_{n-\delta_{ij}} \right) A_n \right) \setminus \left( A_n \left( \bigcup_{i=0}^{k-1} \prod_{j=0}^{k-1} A_{n-\delta_{ij}} \right) A_n \right). \quad (7)$$

From (6) we see that  $x$  must be an element in

$$(A_n^{k+1} A_{n-1}) \setminus \left( A_n \left( \bigcup_{i=0}^{k-1} \prod_{j=0}^{k-1} A_{n-\delta_{ij}} \right) A_n \right).$$

This gives that  $x$  must end with a word in  $(A_n A_{n-1}) \setminus ((A_n A_n)[1 + l_n - l_{n-1}, 2l_n])$ . But then  $x$  cannot be an element of (7), since all its elements ends with a word in  $((A_n A_n)[1 + l_n - l_{n-1}, 2l_n])$ . A contradiction, hence must we have equality in (5).  $\square$

Let us turn to proving a recursion for the size of the  $A_n$  sets. The special case  $k = m = 1$  of (9) was already considered by Godrèche and Luck in [5], see also Nilsson [11].

**Proposition 5.** *For a fixed  $m \geq 1$  we have for  $k \geq 1$  and  $n \geq 2$*

$$\left| \bigcup_{i=0}^k \prod_{j=0}^k A_{n-\delta_{ij}} \right| = \frac{m(n-2) + k + 1}{m(n-2) + 1} |A_n|^k |A_{n-1}| \quad (8)$$

and in particular we have the special case  $k = m$ ,

$$|A_{n+1}| = \frac{m(n-1) + 1}{m(n-2) + 1} |A_n|^m |A_{n-1}|. \quad (9)$$

*Proof.* We give a proof by induction on  $n$ . For the basis case  $n = 2$  we have

$$\left| \bigcup_{i=0}^k \prod_{j=0}^k A_{2-\delta_{ij}} \right| = k+1 = \frac{m(2-2) + k+1}{m(2-2) + 1} |A_2| |A_1|$$

for  $1 \leq k$ . Now assume for induction that (8) holds for  $2 \leq n \leq p$ . Then for the induction step  $n = p+1$  we give a proof by induction on  $k$ . For the basis case  $k = 1$  we have by the recursive definition of the  $A_n$  sets, the induction assumption and by (9)

$$\begin{aligned} \left| \bigcup_{i=0}^1 \prod_{j=0}^1 A_{p+1-\delta_{ij}} \right| &= |A_p A_{p+1} \cup A_{p+1} A_p| \\ &= \left| \left( A_p \left( \bigcup_{i=0}^m \prod_{j=0}^m A_{p-\delta_{ij}} \right) \right) \cup \left( \left( \bigcup_{i=0}^m \prod_{j=0}^m A_{p-\delta_{ij}} \right) A_p \right) \right| \\ &= \left| \bigcup_{i=0}^{m+1} \prod_{j=0}^{m+1} A_{p-\delta_{ij}} \right| \\ &= \frac{m(p-2) + m+2}{m(p-2) + 1} |A_p|^{m+1} |A_{p-1}| \\ &= \frac{m(p-2) + m+2}{m(p-2) + 1} \cdot \frac{m(p-2) + 1}{m(p-1) + 1} |A_{p+1}| |A_p| \\ &= \frac{m(p-1) + 1 + 1}{m(p-1) + 1} |A_{p+1}| |A_p|, \end{aligned}$$

which completes the first basis step. For the second basis step  $k = 2$  we similarly have

$$\begin{aligned} \left| \bigcup_{i=0}^2 \prod_{j=0}^2 A_{p+1-\delta_{ij}} \right| &= \\ &= \left| \left( A_{p+1} \left( \bigcup_{i=0}^1 \prod_{j=0}^1 A_{p+1-\delta_{ij}} \right) \right) \cup \left( \left( \bigcup_{i=0}^1 \prod_{j=0}^1 A_{p+1-\delta_{ij}} \right) A_{p+1} \right) \right| \\ &= 2 |A_{p+1}| \left| \bigcup_{i=0}^1 \prod_{j=0}^1 A_{p+1-\delta_{ij}} \right| - |A_{p+1}|^2 |A_p| \\ &= 2 \cdot \frac{m(p-1) + 1 + 1}{m(p-1) + 1} |A_{p+1}|^2 |A_p| - |A_{p+1}|^2 |A_p| \end{aligned}$$

$$\begin{aligned}
&= \left( 2 \cdot \frac{m(p-1) + 1 + 1}{m(p-1) + 1} - 1 \right) |A_{p+1}|^2 |A_p| \\
&= \frac{m(p-1) + 2 + 1}{m(p-1) + 1} |A_{p+1}|^2 |A_p|,
\end{aligned}$$

which completes the second basis step on  $k$ . Now assume for induction that (8) holds for  $1 \leq k \leq q$ . Then we have for the induction step,  $k = q + 1$  by using Proposition 4 and the induction assumption

$$\begin{aligned}
&\left| \bigcup_{i=0}^{q+1} \prod_{j=0}^{q+1} A_{p+1-\delta_{ij}} \right| = \\
&= \left| \left( A_{p+1} \left( \bigcup_{i=0}^q \prod_{j=0}^q A_{p+1-\delta_{ij}} \right) \right) \cup \left( \left( \bigcup_{i=0}^q \prod_{j=0}^q A_{p+1-\delta_{ij}} \right) A_{p+1} \right) \right| \\
&= 2|A_{p+1}| \left| \bigcup_{i=0}^q \prod_{j=0}^q A_{p+1-\delta_{ij}} \right| - |A_{p+1}|^2 \left| \bigcup_{i=0}^{q-1} \prod_{j=0}^{q-1} A_{p+1-\delta_{ij}} \right| \\
&= 2|A_{p+1}| \frac{m(p-1) + q + 1}{m(p-1) + 1} |A_{p+1}|^q |A_p| \\
&\quad - |A_{p+1}|^2 \frac{m(p-1) + q}{m(p-1) + 1} |A_{p+1}|^{q-1} |A_p| \\
&= \left( 2 \cdot \frac{m(p-1) + q + 1}{m(p-1) + 1} - \frac{m(p-1) + q}{m(p-1) + 1} \right) |A_{p+1}|^{q+1} |A_p| \\
&= \frac{m(p-1) + q + 1 + 1}{m(p-1) + 1} |A_{p+1}|^{q+1} |A_p|,
\end{aligned}$$

which completes the induction on  $k$ , and thereby also the induction on  $n$ .  $\square$

We can now give the explicit way of calculating the size of the set  $A_n$ , by simply unwinding the recursion given in (9).

**Proposition 6.** *For a fixed  $m \geq 1$  we have for  $n \geq 3$*

$$|A_n| = \prod_{i=2}^{n-1} (m(n-i) + 1)^{d_{i-1}} \quad (10)$$

where  $d_i = m \cdot d_{i-1} + d_{i-2}$  with  $d_1 = 1$  and  $d_2 = m - 1$ .

*Proof.* We give a proof by induction on  $n$ . For the basis case  $n = 3$  we have by the recursion (9)

$$|A_3| = \frac{m(3-2) + 1}{m(3-3) + 1} |A_2|^m |A_1| = m(3-2) + 1,$$

and similarly for the basis case  $n = 4$

$$\begin{aligned}
|A_4| &= \frac{m(4-2)+1}{m(4-3)+1} |A_3|^m |A_2| \\
&= \frac{m(4-2)+1}{m(4-3)+1} (m(3-2)+1)^m \\
&= (m(4-2)+1)(m(4-3)+1)^{m-1}.
\end{aligned}$$

Now assume that (10) holds for  $3 \leq n \leq p$ . Then we have for the induction step,  $n = p+1$ , by (9) and the induction assumption

$$\begin{aligned}
|A_{p+1}| &= \\
&= \frac{m(p-1)+1}{m(p-2)+1} |A_p|^m |A_{p-1}| \\
&= \frac{m(p-1)+1}{m(p-2)+1} \left( \prod_{i=2}^{p-1} (m(p-i)+1)^{d_{i-1}} \right)^m \left( \prod_{i=2}^{p-2} (m(p-1-i)+1)^{d_{i-1}} \right) \\
&= \frac{m(p-1)+1}{m(p-2)+1} \left( \prod_{i=2}^{p-1} (m(p-i)+1)^{d_{i-1}} \right)^m \left( \prod_{i=3}^{p-1} (m(p-i)+1)^{d_{i-2}} \right) \\
&= (m(p-1)+1)(m(p-2)+1)^{m-1} \left( \prod_{i=3}^{p-1} (m(p-i)+1)^{m \cdot d_{i-1} + d_{i-2}} \right) \\
&= (m(p-1)+1)^{d_1} (m(p-2)+1)^{d_2} \left( \prod_{i=3}^{p-1} (m(p-i)+1)^{d_i} \right) \\
&= (m(p-1)+1)^{d_1} (m(p-2)+1)^{d_2} \left( \prod_{i=4}^p (m(p+1-i)+1)^{d_{i-1}} \right) \\
&= \prod_{i=2}^{(p+1)-1} (m(p+1-i)+1)^{d_{i-1}},
\end{aligned}$$

which completes the induction.  $\square$

### 3 PREFIX SETS

In this section we study some basic structures of the sets of prefixes of words from  $A_n$ .

**Proposition 7.** *For a fixed  $m \geq 1$  and any  $n \geq 3$  and  $k \geq 0$  we have*

$$A_n[1, l_n - 1] = A_{n+k}[1, l_n - 1]. \quad (11)$$



*Proof.* It is by definition clear that we have the inclusion  $A_n[1, l_n - 1] \subseteq A_{n+k}[1, l_n - 1]$  for  $n \geq 3$  and  $k \geq 0$ . For the reversed inclusion we give a proof by induction on  $k$  and on  $n$ . Let us first, in the basis case  $k = 1$ , give two identities that we shall make use of later on in the proof. We have

$$\begin{aligned}
A_{n+1}[1, l_n - 1] &= \left( \bigcup_{i=0}^m \prod_{j=0}^m A_{n-\delta_{ij}} \right) [1, l_n - 1] \\
&= \left( (A_{n-1} A_n^m) \bigcup \left( \bigcup_{i=1}^m \prod_{j=0}^m A_{n-\delta_{ij}} \right) \right) [1, l_n - 1] \\
&= ((A_{n-1} A_n)[1, l_n - 1]) \bigcup (A_n[1, l_n - 1]), \tag{12}
\end{aligned}$$

and secondly

$$\begin{aligned}
A_n A_{n+1} &= A_n \bigcup_{i=0}^m \prod_{j=0}^m A_{n-\delta_{ij}} \\
&= \bigcup_{i=0}^m \prod_{j=-1}^m A_{n-\delta_{ij}} \\
&= \left( \bigcup_{i=0}^{m-1} \prod_{j=-1}^m A_{n-\delta_{ij}} \right) \bigcup (A_n^{m+1} A_{n-1}) \\
&= \left( \left( \bigcup_{i=0}^{m-1} \prod_{j=-1}^{m-1} A_{n-\delta_{ij}} \right) A_n \right) \bigcup (A_n^{m+1} A_{n-1}) \\
&\subseteq \left( \left( \bigcup_{i=-1}^{m-1} \prod_{j=-1}^{m-1} A_{n-\delta_{ij}} \right) A_n \right) \bigcup (A_n^{m+1} A_{n-1}) \\
&= (A_{n+1} A_n) \bigcup (A_n^{m+1} A_{n-1}). \tag{13}
\end{aligned}$$

Now let us turn to the first basis case  $n = 3$  in the induction on  $n$ . By the identity (12) we have

$$\begin{aligned}
A_4[1, l_3 - 1] &= ((A_2 A_3)[1, l_3 - 1]) \bigcup (A_3[1, l_3 - 1]) \\
&= ((\{1\} A_3)[1, l_3 - 1]) \bigcup (A_3[1, l_3 - 1]) \\
&\subseteq (A_3[1, l_3 - 1]) \bigcup (A_3[1, l_3 - 1]) \\
&= A_3[1, l_3 - 1],
\end{aligned}$$

which concludes the case  $n = 3$ . For the second basis case  $n = 4$  we have by (12)

$$A_5[1, l_4 - 1] = ((A_3 A_4)[1, l_4 - 1]) \bigcup (A_4[1, l_4 - 1]). \quad (14)$$

The next step is to simplify the right hand side of (14). By (13) we have

$$\begin{aligned} (A_3 A_4)[1, l_4 - 1] &\subseteq ((A_4 A_3) \bigcup (A_3^{m+1} A_2))[1, l_4 - 1] \\ &= (A_4[1, l_4 - 1]) \bigcup ((A_3^{m+1} A_2)[1, l_4 - 1]) \\ &= (A_4[1, l_4 - 1]) \bigcup A_3^m \\ &= A_4[1, l_4 - 1], \end{aligned}$$

which, combined by (14) concludes the case. Now assume for induction on  $n$  that (11) holds for  $k = 1$  and  $3 \leq n \leq p$ . For the induction step  $n = p + 1$ , we have by (12)

$$A_{p+2}[1, l_{p+1} - 1] = ((A_p A_{p+1})[1, l_{p+1} - 1]) \bigcup (A_{p+1}[1, l_{p+1} - 1]). \quad (15)$$

As above, we aim to simplify (15). The identity (13) and the induction assumption gives

$$\begin{aligned} (A_p A_{p+1})[1, l_{p+1} - 1] &\subseteq \\ &\subseteq ((A_{p+1} A_p) \bigcup (A_p^{m+1} A_{p-1}))[1, l_{p+1} - 1] \\ &= (A_{p+1}[1, l_{p+1} - 1]) \bigcup (A_p^{m+1}[1, l_{p+1} - 1]) \\ &= (A_{p+1}[1, l_{p+1} - 1]) \bigcup (A_p^m (A_p[1, l_{p-1} - 1])) \\ &= (A_{p+1}[1, l_{p+1} - 1]) \bigcup (A_p^m (A_{p-1}[1, l_{p-1} - 1])) \\ &= (A_{p+1}[1, l_{p+1} - 1]) \bigcup ((A_p^m A_{p-1})[1, l_{p+1} - 1]) \\ &= A_{p+1}[1, l_{p+1} - 1]. \end{aligned} \quad (16)$$

Combining (15) with (16) now gives

$$A_{p+2}[1, l_{p+1} - 1] = A_{p+1}[1, l_{p+1} - 1],$$

which concludes the induction on  $n$  and the basis case  $k = 1$  for the induction on  $k$ . Now assume for induction on  $k$  that (11) holds for  $1 \leq k \leq q$  and

$n \geq 3$ . Then we have for the induction step,  $k = q+1$ , by using the induction assumption twice

$$\begin{aligned}
A_{n+q+1}[1, l_n - 1] &= (A_{n+q+1}[1, l_{n+1} - 1])[1, l_n - 1] \\
&= (A_{(n+1)+q}[1, l_{n+1} - 1])[1, l_n - 1] \\
&= (A_{n+1}[1, l_{n+1} - 1])[1, l_n - 1] \\
&= A_{n+1}[1, l_n - 1] \\
&= A_n[1, l_n - 1],
\end{aligned}$$

which concludes the induction and the proof.  $\square$

The symmetry in the recursive definition of the  $A_n$  sets (1) gives that there clearly is a symmetric analogue of Proposition 7, that is the result holds for sets of suffixes. Next let us construct, by the help of Proposition 7, a superset to  $A_n$  avoiding all the unions in the definition (1).

**Proposition 8.** *Let*

$$\begin{aligned}
P_{n-1} &= (A_{n-1}[1, l_{n-1} - 1])\{0, 1\}, \\
C_{n-1} &= \{0, 1\}(A_{n-1}[2, l_{n-1} - 1])\{0, 1\}, \\
S_{n-2} &= \{0, 1\}(A_{n-2}[2, l_{n-2}]).
\end{aligned}$$

*Then for  $n \geq 4$  we have the inclusion*

$$A_n \subseteq P_{n-1}C_{n-1}^{m-1}S_{n-2}.$$

*Proof.* Let  $B_n = P_{n-1}C_{n-1}^{m-1}S_{n-2}$  and let  $T_i = A_{n-1}^i A_{n-2} A_{n-1}^{m-i}$  for  $0 \leq i \leq m$ . We have to show that  $T_i \subseteq B_n$  for all  $0 \leq i \leq m$ . The case  $i = m$  is direct by the construction, therefore we may assume  $i < m$ . By Proposition 7 we have

$$T_i[1, (i+1)l_{n-1}] \subseteq A_n[1, (i+1)l_{n-1}] \subseteq B_n[1, (i+1)l_{n-1}]. \quad (17)$$

Furthermore, we have from Proposition 7

$$\begin{aligned}
A_{n-1}[1 + l_{n-1} - l_{n-2}, l_{n-1}] &\subseteq \{0, 1\}(A_{n-1}[2 + l_{n-1} - l_{n-2}, l_{n-1}]) \\
&= \{0, 1\}(A_{n-2}[2, l_{n-2}]) \quad (18)
\end{aligned}$$

$$\subseteq \{0, 1\}(A_{n-1}[2, l_{n-2}]). \quad (19)$$

Now combining (17) and (19) gives

$$T_i[1, (i+1)l_{n-1} + l_{n-2}] \subseteq B_n[1, (i+1)l_{n-1} + l_{n-2}].$$

and in particular we have

$$\begin{aligned} T_i[1 + i \cdot l_{n-1} + l_{n-2}, (i+1)l_{n-1} + l_{n-2}] &\subseteq \\ &\subseteq B_n[1 + i \cdot l_{n-1} + l_{n-2}, (i+1)l_{n-1} + l_{n-2}]. \end{aligned} \quad (20)$$

Applying the inclusion (20)  $m-1$  times and combining it the last time with (18) concludes the proof.  $\square$

#### 4 FACTOR SETS

The aim of this section is to give a finite method for finding the factor set  $F_n$ . In [11] we find a finite method for finding the factor set of  $A_n$  in the case  $m = 1$ .

**Proposition 9.** *For a fixed  $m \geq 1$  and all  $n \geq 3$  and  $k \geq 1$  we have*

$$F(A_{n+k}, l_n) = F(A_n^2, l_n). \quad (21)$$

*Proof.* We give a proof by induction on  $k$ . For the basis case  $k = 1$  we look first at the special case  $m = 1$ . It is clear that  $F(A_{n+1}, l_n) \subseteq F(A_n^2, l_n)$ . For the reversed inclusion we have for  $1 \leq i < l_{n-1}$  by Proposition 7,

$$A_n^2[i, i-1 + l_n] = (A_n A_{n-1})[i, i-1 + l_n] \subseteq A_{n+1}[i, i-1 + l_n].$$

The remaining cases of  $i$  follow by symmetry.

For the case  $m \geq 2$ , it is clear that  $F(A_n^2, l_n) \subseteq F(A_{n+1}, l_n)$ . For the reversed inclusion, we have by the recursive definition of the  $A_n$  sets that a sub-word  $w$  of length  $l_n$  of a word in  $A_{n+1}$  is an element of the set

$$F(A_n^2, l_n) \bigcup F(A_{n-1}A_n, l_n) \bigcup F(A_nA_{n-1}, l_n) \bigcup F(A_nA_{n-1}A_n, l_n). \quad (22)$$

For all but the right-most set in (22) it is clear that they are subsets of  $F(A_n^2, l_n)$ . To deal with the remaining case assume  $l_{n-1} < i \leq \lceil \frac{1}{2}(l_n + l_{n-1}) \rceil$ . This implies that

$$i-1 + l_n \leq l_n + m \cdot l_{n-1}, \quad (23)$$

as  $n \geq 3$  and  $m \geq 2$ . By the recursive definition of the  $A_n$  sets (1) and the bound (23) on  $i$  we obtain

$$\begin{aligned} (A_n A_{n-1} A_n)[i, i-1 + l_n] &= \\ &= \left( A_n A_{n-1} \bigcup_{i=0}^m \prod_{j=0}^m A_{n-1-\delta_{ij}} \right) [i, i-1 + l_n] \end{aligned}$$

$$\begin{aligned}
&= \left( A_n \bigcup_{i=0}^m \prod_{j=-1}^m A_{n-1-\delta_{ij}} \right) [i, i-1+l_n] \\
&= \left( \left( A_n \bigcup_{i=0}^{m-1} \prod_{j=-1}^m A_{n-1-\delta_{ij}} \right) \bigcup (A_n A_{n-1}^{m+1} A_{n-2}) \right) [i, i-1+l_n] \\
&\subseteq \left( \left( A_n \bigcup_{i=-1}^{m-1} \prod_{j=-1}^m A_{n-1-\delta_{ij}} \right) \bigcup (A_n A_{n-1}^{m+1} A_{n-2}) \right) [i, i-1+l_n] \\
&= \left( (A_n A_n A_{n-1}) \bigcup (A_n A_{n-1}^{m+1} A_{n-2}) \right) [i, i-1+l_n] \\
&= (A_n A_n) [i, i-1+l_n] \bigcup (A_n A_{n-1}^{m+1}) [i, i-1+l_n] \\
&= (A_n A_n) [i, i-1+l_n] \bigcup (A_n A_{n-1}^m A_{n-2}) [i, i-1+l_n] \\
&= (A_n A_n) [i, i-1+l_n]
\end{aligned}$$

Therefore we see that (22) is a subset to  $F(A_n^2, l_n)$  when  $l_{n-1} < i \leq \lceil \frac{1}{2}(l_n + l_{n-1}) \rceil$ . The remaining cases of  $i$  follows by symmetry.

Now assume for induction that (21) holds for  $1 \leq k \leq q$ . Then for  $k = q+1$  we have by using the basis case and by the induction assumption

$$\begin{aligned}
F(A_{n+q+1}, l_n) &= F(F(A_{n+q+1}, l_{n+1}), l_n) \\
&= F(F(A_{n+1+1}, l_{n+1}), l_n) \\
&= F(A_{n+2}, l_n) \\
&= F(A_{n+1}, l_n),
\end{aligned}$$

which completes the induction.  $\square$

**Proposition 10.** *For a fixed  $m \geq 1$  and all  $n \geq 3$  we have*

$$F_n = F(A_n^2, l_n). \quad (24)$$

*Proof.* It is clear that  $F(A_{n+1}, l_n) \subseteq F_n$ . For the reverse inclusion let  $x \in F_n$ . Then there are words  $u$  and  $v$  such that  $uxv \in A_{n+k}$  for some integer  $k \geq 1$ . That is  $x \in F(A_{n+k}, l_n)$ . Proposition 9 now gives  $F(A_{n+k}, l_n) = F(A_{n+1}, l_n) = F(A_n^2, l_n)$  completing the proof.  $\square$

## 5 UPPER BOUND

The aim of this section is to give an upper bound of the size of  $|F_n|$  in terms of  $|A_n|$ .

**Proposition 11.** *For a fixed  $m \geq 1$  and for any  $n \geq 3$  we have for all  $1 \leq k \leq l_n - 1$*

$$|A_n[1, k]| \cdot |A_n[k + 1, l_n]| \leq 4^{mn} |A_n|. \quad (25)$$

*Proof.* By symmetry we only need to consider the case with  $1 \leq k \leq \lfloor \frac{1}{2} l_n \rfloor$ . We give a proof by induction on  $n$ . For the basis case  $n = 3$  we have

$$|A_3[1, k]| \cdot |A_3[k + 1, l_3]| \leq |A_3|^2 = (m + 1)^2 \leq 4^{3m} (m + 1) = 4^{3m} |A_3|.$$

Now assume for induction that (25) holds for  $3 \leq n \leq p$ . For the induction step  $n = p + 1$  let  $0 \leq i \leq m - 1$  be such that

$$i \cdot l_p \leq k < (i + 1) \cdot l_p.$$

If  $i = 0$  we have by Proposition 8, the induction assumption and (9), when reusing the notation from Proposition 8,

$$\begin{aligned} & |A_{p+1}[1, k]| \cdot |A_{p+1}[k + 1, l_{p+1}]| \leq \\ & \leq |(P_p C_p^{m-1} S_{p-1})[1, k]| \cdot |(P_p C_p^{m-1} S_{p-1})[k + 1, l_{p+1}]| \\ & = |P_p[1, k]| \cdot |P_p[k + 1, l_p]| |C_p|^{m-1} |S_{p-1}| \\ & = |A_p[1, k]| \cdot |A_p[k + 1, l_p]| |A_p|^{m-1} 4^m |A_{p-1}| \\ & \leq 4^{mp} |A_p| 4^m |A_p|^{m-1} |A_{p-1}| \\ & = 4^{m(p+1)} |A_p|^m |A_{p-1}| \\ & \leq 4^{m(p+1)} |A_{p+1}|, \end{aligned}$$

completing the induction in this case. Similarly, for the case  $i \geq 1$  we have

$$\begin{aligned} & |A_{p+1}[1, k]| \cdot |A_{p+1}[k + 1, l_{p+1}]| \leq \\ & \leq |(P_p C_p^{m-1} S_{p-1})[1, k]| \cdot |(P_p C_p^{m-1} S_{p-1})[k + 1, l_{p+1}]| \\ & = |(P_p C_p^i)[1, k]| \cdot |C_p^{m-i}[k + 1 - i \cdot l_p, (m - i)l_p]| |S_{p-1}| \\ & = 4^i |A_p|^i |A_p[1, k - i \cdot l_p]| \cdot |A_p[k + 1 - i \cdot l_p, l_p]| |A_p|^{m-i-1} 4^{m-i} |A_{p-1}| \\ & = 4^m |A_p|^{m-1} |A_p[1, k - i \cdot l_p]| \cdot |A_p[k + 1 - i \cdot l_p, l_p]| |A_{p-1}| \\ & \leq 4^m |A_p|^{m-1} 4^{mp} |A_p| |A_{p-1}| \\ & = 4^{m(p+1)} |A_p|^m |A_{p-1}| \\ & \leq 4^{m(p+1)} |A_{p+1}|, \end{aligned}$$

which completes the case and the proof.  $\square$

The estimate in Proposition 11 is probably far from optimal. Computer calculations lead us to conjecture that we may replace the factor  $4^{mn}$  in (25) by a constant.

We can now give the estimate we set out to find. Again compared to computer calculations our estimate seems to be far from the optimal one, but it will be sufficient for our purpose.

**Proposition 12.** *For a fixed  $m \geq 1$  we have for  $n \geq 3$*

$$|F_n| \leq 4^{mn} \cdot l_n \cdot |A_n|.$$

*Proof.* By Proposition 10 it is enough to estimate the size of  $F(A_n^2, l_n)$ . Let  $x$  be a subword of length  $l_n$  in  $A_n^2$ . That is, there is a  $0 \leq k < l_n$  such that

$$x \in A_n^2[k+1, k+l_n] = (A_n[k+1, l_n])(A_n[1, k]).$$

By Proposition 11 there are at most  $4^{mn}|A_n|$  such words  $x$ . Adding up for all possible values of  $k$  gives the desired result.  $\square$

## 6 PROOF OF THE MAIN RESULT

Finally, we have now gathered enough background to prove our theorem.

*Proof of Theorem 3.* It is already clear that the limit  $\lim_{n \rightarrow \infty} \frac{\log |F_n|}{l_n}$  exists. By Proposition 12 we have  $|A_n| \leq |F_n| \leq 4^{mn} l_n |A_n|$  and by Proposition 2 we have

$$l_n = \frac{1}{\sqrt{m^2+4}} \left( \frac{m + \sqrt{m^2+4}}{2} \right)^n - \frac{1}{\sqrt{m^2+4}} \left( \frac{m - \sqrt{m^2+4}}{2} \right)^n$$

from which it follows that  $l_n$  grows with  $n$  like  $\left( \frac{m + \sqrt{m^2+4}}{2} \right)^n$ . The monotonicity of the logarithm now implies

$$0 \leq \frac{\log |F_n|}{l_n} - \frac{\log |A_n|}{l_n} \leq \frac{\log(4^{mn} l_n)}{l_n} \rightarrow 0$$

when  $n \rightarrow \infty$ , which gives that the limit  $\lim_{n \rightarrow \infty} \frac{\log |A_n|}{l_n}$  exists and equals  $\lim_{n \rightarrow \infty} \frac{\log |F_n|}{l_n}$ .  $\square$

## 7 ACKNOWLEDGEMENT

The author wishes to thank Michael Baake, Peter Zeiner and Markus Moll at Bielefeld University, Germany, for our discussions of the problem and for reading drafts of the manuscript. This work was supported by the German Research Council (DFG), via CRC 701.

## REFERENCES

- [1] Axel F., Dénoyer F., Gazeau J. P.: From quasicrystals to more complex systems. EDP Sciences, Les Ulis, and Springer, Berlin (2000).
- [2] Baake M., Grimm U.: Kinematic diffraction from a mathematical viewpoint. *Zeitschrift für Kristallographie* 226 (2011) 711–725.
- [3] Baake M., Moody R. V.: Diffractive point sets with entropy. *J. Phys. A: Math. Gen.* 31 9023 doi: 10.1088/0305-4470/31/45/003
- [4] Fekete M.: Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten. *Mathematische Zeitschrift*, no. 17 228–249 (1923).
- [5] Godrèche C., Luck J. M.: Quasiperiodicity and randomness in tilings of the plane. *J. of Stat Phys*, Vol. 55, Nos. 1/2 (1989), 1–28.
- [6] Lagarias J.: Meyer’s concept of quasicrystal and quasiregular sets. *Comm. Math. Phys.* 179 (1996), no. 2, 365–376.
- [7] Lütkehölder C.: Diffraction stochastischer Fibonacci-Mengen. Diplomarbeit, Fakultät für Mathematik, Universität Bielefeld, (2010).
- [8] Meyer Y.: Algebraic numbers and harmonic analysis. Amsterdam North-Holland Publ. Co., (1972).
- [9] Moody R. V.: Meyer sets and their duals. 403–441, *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, 489, Kluwer Acad. Publ., Dordrecht, (1997).
- [10] Moody R. V.: Model sets: A Survey. In [1] (2000) 145–166.
- [11] Nilsson J.: On the entropy of random Fibonacci words. arXiv:1001.3513.
- [12] Strungaru N.: Almost periodic measures and long-range order in Meyer sets. *Discrete Comput. Geom.* 33 (2005), no. 3, 483–505.